

PATENT  
81940.0052

Express Mail Label No. EV 325 216 978 US

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

Seiji KANEKO et al.

Serial No: Not assigned

Filed: July 28, 2003

For: Disk Array Device, Method for Controlling  
Disk Array Device and Storage System

Art Unit: Not assigned

Examiner: Not assigned

TRANSMITTAL OF PRIORITY DOCUMENT

Mail Stop PATENT APPLICATION

Commissioner for Patents

P.O. Box 1450

Alexandria, VA 22313-1450

Dear Sir:

Enclosed herewith is a certified copy of Japanese patent application No. 2002-303524 which was filed October 17, 2002, from which priority is claimed under 35 U.S.C. § 119 and Rule 55.

Acknowledgment of the priority document(s) is respectfully requested to ensure that the subject information appears on the printed patent.

Respectfully submitted,

HOGAN & HARTSON L.L.P.

Date: July 28, 2003

By: 

Anthony J. Orler

Registration No. 41,232

Attorney for Applicant(s)

500 South Grand Avenue, Suite 1900  
Los Angeles, California 90071  
Telephone: 213-337-6700  
Facsimile: 213-337-6701

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日  
Date of Application:

2002年10月17日

出 願 番 号  
Application Number:

特願2002-303524

[ ST.10/C ]:

[ JP2002-303524 ]

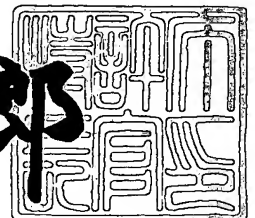
出 願 人  
Applicant(s):

株式会社日立製作所

2003年 4月25日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

太田 信一郎



出証番号 出証特2003-3030709

【書類名】 特許願

【整理番号】 HI020537

【提出日】 平成14年10月17日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 3/06

【発明者】

    【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

    【氏名】 金子 誠司

【発明者】

    【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

    【氏名】 金井 宏樹

【特許出願人】

    【識別番号】 000005108

    【氏名又は名称】 株式会社日立製作所

【代理人】

    【識別番号】 100071283

    【弁理士】

    【氏名又は名称】 一色 健輔

【選任した代理人】

    【識別番号】 100084906

    【弁理士】

    【氏名又は名称】 原島 典孝

【選任した代理人】

    【識別番号】 100098523

    【弁理士】

    【氏名又は名称】 黒川 恵

【選任した代理人】

【識別番号】 100112748

【弁理士】

【氏名又は名称】 吉田 浩二

【選任した代理人】

【識別番号】 100110009

【弁理士】

【氏名又は名称】 青木 康

【手数料の表示】

【予納台帳番号】 011785

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ディスクアレイ装置、ディスクアレイ装置の制御方法、及びストレージシステム

【特許請求の範囲】

【請求項 1】 外部装置からのデータ入出力要求を受信する複数の入出力チャネルと、

入出力チャネルごとに用意され、それぞれ対応する入出力チャネルに接続する複数のキャッシュメモリと、

ディスクドライブと、

ディスクドライブに対してデータ入出力を行うディスク制御手段と、

入出力チャネルとディスク制御手段とを通信可能に接続する通信手段と、  
を備えることを特徴とするディスクアレイ装置。

【請求項 2】 外部装置からのデータ入出力要求を受信する複数の入出力チャネルと、

入出力チャネルごとに用意され、それぞれ対応する入出力チャネルに接続する複数のキャッシュメモリと、

ディスクドライブと、

ディスクドライブに対してデータ入出力を行うディスク制御手段と、

入出力チャネルとディスク制御手段とを通信可能に接続する通信手段と、

各キャッシュメモリに記憶しているデータに関する整合性を確保するための整合性確保処理を実行する手段と、

前記データ入出力要求を受信した場合に、そのデータ入出力要求の内容に応じて、そのデータ入出力要求に対するホストコンピュータへの応答処理と前記整合性確保処理との実行順を制御する手段と、

を備えることを特徴とするディスクアレイ装置。

【請求項 3】 請求項 2 に記載のディスクアレイ装置において、  
前記整合性確保処理には、

ある前記キャッシュメモリに記憶しているデータについての更新が行われる場合に、そのデータが記憶されている前記ディスクドライブの記憶領域と同一の記

憶領域に記憶される他の前記キャッシュメモリに記憶しているデータを無効化する処理が含まれること、

を特徴とする請求項 2 に記載のディスクアレイ装置。

【請求項 4】 請求項 2 に記載のディスクアレイ装置において、  
前記整合性確保処理には、

ある前記キャッシュメモリに記憶しているデータについての更新が行われる場合に、当該キャッシュメモリの当該データを更新するとともに、そのデータが記憶されている前記ディスクドライブの記憶領域と同一の記憶領域に記憶される他の前記キャッシュメモリに記憶されているデータについても更新する処理が含まれること、

を特徴とする請求項 2 に記載のディスクアレイ装置。

【請求項 5】 請求項 2 に記載のディスクアレイ装置において、  
前記ディスクドライブの記憶領域上には区画設定された論理的な記憶領域である複数の論理ボリュームが設定され、

前記データ入出力要求には当該データ入出力要求の処理対象となる論理ボリュームの識別子が設定され、

前記実行順の制御を、前記データ入出力要求に設定されている前記論理ボリュームの識別子に応じて行う手段を備えること、

を特徴とする請求項 2 に記載のディスクアレイ装置。

【請求項 6】 請求項 2 に記載のディスクアレイ装置において、  
他のディスクアレイ装置と接続する手段と、

データ入出力要求としてデータ書き込み要求を受信した場合に、この要求で指定されるデータを自身のディスクドライブに書き込むとともにそのデータについての書き込み要求を前記他のディスクアレイ装置に送信する手段を備え、

前記入出力チャネルは、データ書き込み要求を受信した場合に、

自身のディスクドライブにデータを書き込む動作モードと、

他のディスクアレイ装置に前記書き込み要求を送信する動作モードと、

のいずれかのモードで動作し、

前記実行順の制御を、前記外部装置からのデータ入出力要求が入力される前記

入出力チャンネルが、前記いずれの動作モードで動作しているかに応じて行う手段と、を備えること、

を特徴とする請求項 2 に記載のディスクアレイ装置。

【請求項 7】 外部装置からのデータ入出力要求を受信する複数の入出力チャンネルと、

入出力チャンネルごとに用意され、それぞれ対応する入出力チャンネルに接続する複数のキャッシュメモリと、

ディスクドライブと、

ディスクドライブに対してデータ入出力を行うディスク制御手段と、

入出力チャンネルとディスク制御手段とを通信可能に接続する通信手段と、

各キャッシュメモリに記憶しているデータに関する整合性を確保するための整合性確保処理を実行する手段と、

を備えるディスクアレイ装置の制御方法において、

前記データ入出力要求を受信した場合に、そのデータ入出力要求の内容に応じて、そのデータ入出力要求に対するホストコンピュータへの応答処理と前記整合性確保処理との実行順を制御すること、

を特徴とするディスクアレイ装置の制御方法。

【請求項 8】 請求項 7 に記載のディスクアレイ装置の制御方法において、前記整合性確保処理には、

ある前記キャッシュメモリに記憶しているデータについての更新が行われる場合に、そのデータが記憶されている前記ディスクドライブの記憶領域と同一の記憶領域に記憶される他の前記キャッシュメモリに記憶しているデータを無効化する処理が含まれること、

を特徴とする請求項 7 に記載のディスクアレイ装置の制御方法。

【請求項 9】 請求項 7 に記載のディスクアレイ装置の制御方法において、前記整合性確保処理には、

ある前記キャッシュメモリに記憶しているデータについての更新が行われる場合に、当該キャッシュメモリの当該データを更新するとともに、そのデータが記憶されている前記ディスクドライブの記憶領域と同一の記憶領域に記憶される他

の前記キャッシュメモリに記憶されているデータについても更新する処理が含まれること、

を特徴とする請求項 7 に記載のディスクアレイ装置の制御方法。

【請求項 1 0】 請求項 7 に記載のディスクアレイ装置の制御方法において

前記ディスクドライブの記憶領域上には区画設定された論理的な記憶領域である複数の論理ボリュームが設定され、

前記データ入出力要求には当該データ入出力要求の処理対象となる論理ボリュームの識別子が設定され、

前記実行順の制御は、前記データ入出力要求に設定されている前記論理ボリュームの識別子に応じて行われることを特徴とする請求項 7 に記載のディスクアレイ装置の制御方法。

【請求項 1 1】 請求項 7 に記載のディスクアレイ装置の制御方法において

ディスクアレイ装置には他のディスクアレイ装置が接続し、

ディスクアレイ装置は、データ入出力要求としてデータ書き込み要求を受信した場合に、この要求で指定されるデータを自身のディスクドライブに書き込むとともにそのデータについての書き込み要求を前記他のディスクアレイ装置に送信し、

前記入出力チャネルは、データ書き込み要求を受信した場合に、

自身のディスクドライブにデータを書き込む動作モードと、

他のディスクアレイ装置に前記書き込み要求を送信する動作モードと、

のいずれかのモードで動作し、

前記実行順の制御は、前記外部装置からのデータ入出力要求が入力される前記入出力チャネルが、前記いずれの動作モードで動作しているかに応じて行われることを特徴とする請求項 7 に記載のディスクアレイ装置の制御方法。

【請求項 1 2】 外部装置と、外部装置からのデータ入出力要求を受信する複数の入出力チャネルと、入出力チャネルごとに用意され、それぞれ対応する入出力チャネルに接続する複数のキャッシュメモリと、ディスクドライブと、ディ



スクドライブに対してデータ入出力を行うディスク制御手段と、入出力チャネルとディスク制御手段とを通信可能に接続する通信手段と、各キャッシュメモリに記憶しているデータに関する整合性を確保するための整合性確保処理を実行するディスクアレイ装置とを備え、

前記ディスクアレイ装置が、前記データ入出力要求を受信した場合に、そのデータ入出力要求の内容に応じて、そのデータ入出力要求に対するホストコンピュータへの応答処理と前記整合性確保処理との実行順を制御する手段を備えること

を特徴とするストレージシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、ディスクアレイ装置、ディスクアレイ装置の制御方法、及びストレージシステムに関する。

【0002】

【従来の技術】

一般的なディスクアレイ装置では、ディスクドライブへのデータの入出力を制御するディスクアダプタと、ホストコンピュータとの接続を制御する入出力チャネルと、キャッシュメモリとがディスクアレイ装置の内部に設けられたネットワークスイッチを介して接続する構成となっている。

【0003】

【特許文献1】

特開 2 0 0 1 - 2 9 0 7 9 0 号公報

【0004】

【発明が解決しようとする課題】

このため、ディスクアレイ装置のネットワークスイッチには、入出力チャネル、ディスクアダプタ、キャッシュメモリの間でのトラフィックが競合し、これによりネットワークスイッチの利用率が高くなり、これがホストコンピュータへの応答速度を低下させる要因となっていた。

本発明はこのような観点からなされたもので、ディスクアレイ装置、ディスクアレイ装置の制御方法、及びストレージシステムを提供することを目的とする。

【 0 0 0 5 】

【課題を解決するための手段】

この目的を達成するための主たる発明であるディスクアレイ装置は、  
外部装置からのデータ入出力要求を受信する複数の入出力チャンネルと、  
入出力チャンネルごとに用意され、それぞれ対応する入出力チャンネルに接続する  
複数のキャッシュメモリと、  
ディスクドライブと、  
ディスクドライブに対してデータ入出力を行うディスク制御手段と、  
入出力チャンネルとディスク制御手段とを通信可能に接続する通信手段と、  
を備えることを特徴とする。

なお、本発明の他の特徴については、本明細書及び添付図面の記載により明らかにする。

【 0 0 0 6 】

【発明の実施の形態】

=== 開示の概要 ===

以下の開示により、少なくとももつぎのことが明らかにされる。

前述の目的に対応して、本発明者らは、上記のように、入出力チャンネルごとにキャッシュメモリを設けることを着想した。図1にそのように構成したディスクアレイ装置1を用いて構成したストレージシステムの概略構成を示す。この構成では、ホストコンピュータからのデータ入出力要求の対象となるデータが、そのデータ入出力要求を受信した入出力チャンネル10a～10dに対応して用意されているキャッシュメモリ11a～11dに存在する場合には、そのデータをホストコンピュータに返すことができるので、キャッシュメモリ11a～11dにデータが存在する場合には、ホストコンピュータに迅速に応答を返すことが可能である。

【 0 0 0 7 】

ところで、図1の構成において、例えば、異なる入出力チャンネル10a～10dで受信したデータ入出力要求により、ディスクドライブの同一の記憶領域に記憶され

ているデータに対してアクセスがあると、ディスクドライブのある記憶領域に記憶されていたデータが複数のキャッシュメモリ11a～11dに同時に記憶されることになる。

【0008】

ここでこの状態において、例えば、ある入出力チャネル10a～10dが受信したデータ入出力要求に起因して前記データに対する更新が行われる場合には、それぞれのキャッシュメモリ11a～11dが記憶しているデータの整合性を保つために、例えば、他のキャッシュメモリ11a～11dに記憶しているデータを無効化してしまう処理（無効化処理）や、他のキャッシュメモリ11a～11dに記憶しているデータを書き換える、といった整合性を確保するための処理を行う必要がある。

【0009】

しかしながら、この整合性を確保するための処理においては、無効化のための通信や各キャッシュメモリ11a～11dへの更新後のデータの通知のための通信などによりネットワークスイッチのトラフィックを増大させ、これによりホストコンピュータに対する応答速度が低下し、入出力チャネル10a～10dごとにキャッシュメモリ11a～11dを設けたことにより得られる効果が希釈することになる。

【0010】

ところで、整合性を確保するための処理は、必ずしも必要でないこともある。例えば、ある記憶領域に記憶されているデータに対するアクセスが、特定の入出力チャネル10a～10dを介してしか行われなような場合である。

【0011】

また、同じ記憶領域に記憶されているデータに対して複数の入出力チャネル10a～10dからのアクセスが想定される場合であっても、ホストコンピュータで動作するソフトウェアの仕様や性質などにより、必ずしも整合性を確保する必要がなかったり、また、必要があったとしても厳密な整合性までは要求されないこともある。

【0012】

そこで、本発明の他の一態様では、外部装置からのデータ入出力要求を受信する複数の入出力チャネルと、入出力チャネルごとに用意され、それぞれ対応する

入出力チャンネルに接続する複数のキャッシュメモリと、ディスクドライブと、ディスクドライブに対してデータ入出力を行うディスク制御手段と、入出力チャンネルとディスク制御手段とを通信可能に接続する通信手段と、各キャッシュメモリに記憶しているデータに関する整合性を確保するための整合性確保処理を実行する手段と、を備えるディスクアレイ装置において、前記データ入出力要求を受信した場合に、そのデータ入出力要求の内容に応じて、そのデータ入出力要求に対するホストコンピュータへの応答処理と前記整合性確保処理との実行順を制御する手段を設けることとした。

## 【 0 0 1 3 】

なお、整合性確保処理には、例えば、ある前記キャッシュメモリに記憶しているデータについての更新が行われる場合に、そのデータが記憶されている前記ディスクドライブの記憶領域と同一の記憶領域に記憶される他の前記キャッシュメモリに記憶しているデータを無効化する処理や、ある前記キャッシュメモリに記憶しているデータについての更新が行われる場合に、当該キャッシュメモリの当該データを更新するとともに、そのデータが記憶されている前記ディスクドライブの記憶領域と同一の記憶領域に記憶される他の前記キャッシュメモリに記憶されているデータについても更新する処理などがある。

## 【 0 0 1 4 】

前述したように、整合性を確保するための処理は、ホストコンピュータで動作するソフトウェアの仕様や性質などにより、必ずしも必要でなかったり、また、必要があるとしても厳密な整合性までは要求されないことがある。そこで、本発明のように、ホストコンピュータから送信されたデータ入出力要求の内容に応じて、そのデータ入出力要求に対するホストコンピュータへの応答処理と整合性確保処理との実行順を制御するようにし、整合性を確保する必要がある場合とそうでない場合とで実行順を制御することで、ディスクアレイ装置を効率よく利用することが可能となる。具体的には、例えば、整合性を確保するための処理が必ずしも必要で無い場合には、ホストコンピュータへの応答処理を行った後に整合性確保処理を実行することで、ホストコンピュータへの応答を迅速に行うようにすることができる。

## 【 0 0 1 5 】

また、ディスクドライブの記憶領域上に論理的な記憶領域である複数の論理ボリュームが設定され、データ入出力要求に当該データ入出力要求の処理対象となる論理ボリュームの識別子が設定されている場合に、前記実行順の制御は、前記データ入出力要求に設定されている前記論理ボリュームの識別子に応じて行われるようにすることもできる。

## 【 0 0 1 6 】

このように論理ボリュームの識別子により実行順を制御することができることで、設定方法のバリエーションが増え、例えば、ホストコンピュータで動作するアプリケーションごとに個別に論理ボリュームを割り当てているような場合には、アプリケーションプログラムの性質に応じて、前記実行順の制御に関する設定を簡単に行うことができる。

## 【 0 0 1 7 】

また、本発明の他の態様では、ディスクアレイ装置には他のディスクアレイ装置が接続しており、ディスクアレイ装置はデータ入出力要求としてデータ書き込み要求を受信した場合に、この要求で指定されるデータを自身のディスクドライブに書き込むとともにそのデータについての書き込み要求を前記他のディスクアレイ装置に送信し、前記入出力チャネルは、データ書き込み要求を受信した場合に、自身のディスクドライブにデータを書き込む動作モード（ノーマルモード）と、他のディスクアレイ装置に前記書き込み要求を送信する動作モード（コピーモード）と、のいずれかのモードで動作し、前記実行順の制御は、前記外部装置からのデータ入出力要求が入力される前記入出力チャネルが、前記いずれの動作モードで動作しているかに応じて行われることを特徴とする。

## 【 0 0 1 8 】

ここで例えば、コピーモードで動作中の入出力チャネルで受信したデータ入出力要求の処理は、ノーマルモードで動作中の入出力チャネルで受信したデータ入出力要求の処理に比べて整合性確保の必要性が少ないような場合には、例えば、ホストコンピュータへの応答処理を行った後に整合性確保処理を実行することで、ホストコンピュータへの応答を迅速に行うようにすることができる。

【 0 0 1 9 】

=== 詳細 ===

## &lt;システム構成&gt;

図 1 のストレージシステムの構成について詳述する。ディスクアレイ装置 1 は、例えば、S A N (Storage Area Network) や L A N (Local Area Network) などの通信手段を介して外部装置であるホストコンピュータに接続している。ディスクアレイ装置 1 は、ホストコンピュータとの間で通信を行う 4 つの入出力チャンネル 10a~10d と、入出力チャンネル 10a~10d ごとに用意されそれぞれ対応する入出力チャンネルに接続するキャッシュメモリ 11a~11d、ディスクドライブ 14a~14d、ディスクドライブ 14a~14d に接続しディスクドライブ 14a~14d へのデータの入出力を制御する 4 つのディスクアダプタ 13a~13d、入出力チャンネル 10a~10d とディスクアダプタ 13a~13d の間を接続する二つのネットワークスイッチ 12a、12b、などを備えて構成される。

【 0 0 2 0 】

なお、図 1 では、ネットワークスイッチ 12a、12b は複数設けられ、入出力チャンネル 10a~10d とディスクアダプタ 13a~13d の間の経路には冗長性を持たせているが、ネットワークスイッチは単数で構成してもよい。4 つの入出力チャンネル 10a~10d には、ホストコンピュータから、データ要求やデータ書き込み要求などのデータ入出力要求が送信されてくる。ディスクアレイ装置 1 は、入出力チャンネル 10a~10d が受信したデータ入出力要求に応じて、ディスクドライブ 14a~14d に対するデータの読み出しや書き込みを行い、必要に応じて、その結果や読み出したデータなどを、ホストコンピュータに送信する。

【 0 0 2 1 】

## &lt;入出力チャンネル&gt;

図 2 は入出力チャンネル 10a~10d の構成の例である。入出力チャンネル 10a~10d はそれぞれプロセッサ 201 を備える。またプロセッサ 201 で実行される制御プログラムや、処理に必要な情報(データ)を記憶するローカルメモリ 202 を備える。さらにホストコンピュータと接続するためのホスト I/F 回路 204 や、ネットワークスイッチ 12a、12b と接続するためのネットワーク接続回路 203、203b を備える。

## 【 0 0 2 2 】

これらは高速な信号線バスであるプロセッサバス205で接続される。さらにキャッシュメモリ11a～11dも、このプロセッサバス205に接続される。この構成によりプロセッサ201はキャッシュメモリ11a～11dに対して読み出しや書き込みを高速に行うことができる。

## 【 0 0 2 3 】

なお、図2ではローカルメモリ202に格納した制御プログラムをプロセッサ201で実行することにより、入出力チャンネル10a～10dの処理を行うことを想定している。制御プログラムの一部または全部の機能をハードウェア化することも考えられる。

## 【 0 0 2 4 】

## &lt; ディスクアダプタ &gt;

図3はディスクアダプタ13a～13dの構成の一例である。ディスクアダプタ13a～13dはそれぞれプロセッサ301を備えている。またプロセッサ301で実行される制御プログラムや、処理に必要な情報(データ)を記憶するために、ローカルメモリ302を備えている。

## 【 0 0 2 5 】

さらにディスクドライブ14a～14dと直接接続するディスクドライブI/F回路304や、ネットワークスイッチ12a, 12bと接続するためのネットワーク接続回路303a, 303bを備える。これらはプロセッサ301と高速な信号線バスであるプロセッサバス305で接続される。

## 【 0 0 2 6 】

なお、図3ではローカルメモリ302に格納した制御プログラムをプロセッサ301で実行することにより、ディスクアダプタの処理を行うことを想定しているが、入出力チャンネルの場合と同様に、制御プログラムの一部または全部の機能をハードウェア化することも考えられる。

## 【 0 0 2 7 】

## &lt; キャッシュメモリ &gt;

図4は、キャッシュメモリ11a～11dの構成の例である。キャッシュメモリは検

索制御部40とメモリ部41とを備え、メモリ部41内に各々のデータを、有効ビットV 401、書き換えビットD 402、データブロック番号403、データ内容404の組で記憶している。

## 【 0 0 2 8 】

入出力チャネル10a～10dは、アドレスバス411によってホストコンピュータに要求されたデータブロック番号を通知し、メモリ部41内でデータブロック番号403が一致し、有効ビットV 401が有効状態（例えば' 1 '）であるデータがあるかどうか検索制御部40に検索させる。

## 【 0 0 2 9 】

検索した結果、データが見つかりと返答信号バス412によってヒット信号が伝達され、見つかったことが報告される。またデータバス413によって、データ内容404が読み出される。データの書き込みの場合には、制御コマンドバス414によって書き込みコマンド信号が通知され、データバス413によって伝達されるデータが当該データ内容404に書き込まれて、書き換えビットD 402がセット状態（例えば' 1 '）になって書き換えが生じたことが記録される。検索した結果、データが見つからない場合には、返答信号バス412によってミス信号が伝達され、見つからなかったことが報告される。

## 【 0 0 3 0 】

ディスクアダプタ13a～13dからキャッシュメモリ上のデータを無効化するための無効化信号が送信されてきた場合には、同様に入出力チャネル10a～10dは、アドレスバス411によって当該データブロック番号を通知し、メモリ部41内でデータブロック番号403が一致し、有効ビットV 401が有効状態（例えば' 1 '）であるデータがあるかどうか検索制御部40に検索させる。

## 【 0 0 3 1 】

検索した結果見つかった場合、書き換えビットD 402の内容も調べる。もし書き換えビットD 402がセット状態である場合には、返答信号バス412によってダーティヒット信号が伝達される。

## 【 0 0 3 2 】

ダーティヒット信号を受け取った入出力チャネル10a～10dは、データ内容404



をデータバス413によって読み出し、無効化信号を送信してきたディスクアダプタ13a~13dに返送する。書き換えビットD 402がセット状態でない場合には、返答信号バス412によってクリーンヒット信号が伝達される。

## 【 0 0 3 3 】

これらの二つの場合において、その後、入出力チャネル10a~10dは、再びアドレスバス411によって当該データブロック番号、制御コマンドバス414によって無効化コマンド信号を通知し、有効ビットV 401を無効状態(例えば'0')としてそのデータの組を無効とする。

## 【 0 0 3 4 】

また検索した結果見つからなかった場合には、返答信号バス412によってミス信号が伝達されて、見つからなかったことが入出力チャネル10a~10dに報告される。これらすべての場合において、最後に入出力チャネル10a~10dは無効化信号を送信してきたディスクアダプタ13a~13dに無効化処理終了信号を返送して、無効化に関する処理が終了したことを伝える。

## 【 0 0 3 5 】

なお、書き換えビットD 402がセット状態になって、ディスクドライブから読み出したデータの書き換えが生じたことが記録された場合、上述の無効化によるほか、適当な時間経過、もしくはキャッシュメモリが溢れた時に、ディスクアダプタ13a~13dに書き換えられたデータを送り、ディスクドライブに書き戻しを行わせたら当該書き換えビットD 402はリセット状態(例えば'0')にし、書き換えが生じたという記録を消去する。

## 【 0 0 3 6 】

なお、図4では、検索制御部40が実際にメモリ部41の検索制御を行うとしたが、メモリ部41だけを備え、検索制御部40は特に備えずに、同等の機能を入出力チャネル10a~10dのプロセッサ201に行わせるようにすることもできる。

## 【 0 0 3 7 】

また図4では、アドレスバス411、返答信号バス412、データバス413、制御コマンドバス414の各種の信号バスがあって、これらが統合されて入出力チャネル10a~10dに接続されるとしたが、単一の信号バスを備え、アドレスバス411、返答

信号バス412、データバス413、制御コマンドバス414の各種の信号バスに相当する信号コマンドが該単一の信号バス上を伝達されるようにすることもできる。

#### 【 0 0 3 8 】

図4の例では、単一のメモリ部41に有効ビットV 401、書き換えビットD 402、データブロック番号403、データ内容404のデータの組がすべて格納され、当該データがあるかどうかを検索する場合には、メモリ部41全体を検索する必要がある。このため、メモリ部41の容量が大きくなってくると検索に時間がかかることが考えられる。

#### 【 0 0 3 9 】

このような場合には図5のようにすることが考えられる。図5では、制御検索部50と、タグ部51、データ部52より構成され、タグ部51には、有効ビットV 501、書き換えビットD 502、データブロック番号503、キャッシュ内データ位置504が格納される。データ部52にはデータ内容505が格納される。

タグ部51において、有効ビットV 501、書き換えビットD 502、データブロック番号503は、図4で示したものと同等の内容を記録する。キャッシュ内データ位置504には、データ部52における当該データの位置が記録される。

データ部52よりデータ内容を取り出す場合には、当該データブロック番号503に対応するキャッシュ内データ位置504より当該データの位置がデータ位置信号515によってデータ部52に伝達され、データバス513に読み出される。

書き込みの時も同様にデータ位置信号515によって指定された位置にデータバス513によって伝達されたデータが書き込まれる。その他のアドレスバス511、制御コマンドバス514、返答信号バス512によって伝達される信号等は、図4の場合と同等である。

#### 【 0 0 4 0 】

このような構成を採ることにより、図5の例では、当該データの有無を検索する際にタグ部51内のみを検索すればよく、データ部512の容量が大きくなっても、検索時間にはあまり影響しないようにすることが可能である。

すなわち、図4の場合には、メモリ部41にデータ内容も一緒に格納されるため、データブロック数が多くなると多数の半導体チップが必要になり、多数の半導

体チップに渡って検索する必要が生じ、検索に要する時間が大きくなる。

これに対し図5の場合には、データ内容を記憶するデータ部が格納される半導体チップと、タグ部を格納する半導体チップは別々になり、タグ部を格納する半導体チップの数は少なくすみ、少ない半導体チップに渡って検索すればよいため、検索に要する時間は小さくなる。

#### 【0041】

なお、図5では、検索制御部50が実際にタグ部51の検索制御を行うとしたが、図4の例と同様に、タグ部51、データ部52だけを備え、検索制御部50は特に備えずに、同等の機能を入出力チャネル10a~10dのプロセッサ201で行わせることも考えられる。

#### 【0042】

また図5では、アドレスバス511、返答信号バス512、データバス513、制御コマンドバス514の各種の信号バスがあって、これらが統合されて入出力チャネル10a~10dに接続されるとしたが、やはり図4の例と同様に、単一の信号バスを備え、アドレスバス511、返答信号バス512、データバス513、制御コマンドバス514の各種の信号バスに相当する信号コマンドが該単一の信号バス上を伝達されるようにしてもよい。

#### 【0043】

##### <データテーブル>

図6は、入出力チャネル10a~10dにおいて、ホストコンピュータから要求されたデータブロック番号のデータが自分に接続されたキャッシュメモリ11a~11dに存在しなかったとき、どのディスクアダプタ13a~13dに当該データを要求すればよいか記録した各入出力チャネルに設けられたデータテーブルの形式の例である。

#### 【0044】

データブロック番号とディスクアダプタ番号が組みとなって記録され、当該入出力チャネル10a~10dにそれぞれ接続されたホストコンピュータとの間で予め取り決めた分だけのデータブロック番号についての組みが存在する。

また、ある個数(例えば16個)の連続したデータブロック番号を同一のディス

クアダプタに接続されたディスクドライブに格納することとして、その個数分(例えば16個)まとめ、先頭のディスクドライブ番号とディスクアダプタ番号の組みを作ってもよい。このようにすると、例えば16分の1にデータテーブルの大きさを縮小することができる。

なお、このようなデータテーブルは、例えば図2におけるローカルメモリ202に格納され、プロセッサ201によって必要に応じて読み書きされる。

#### 【0045】

図7は、ディスクアダプタ13a~13dにおいて、入出力チャネル10a~10dから要求されたデータブロック番号のデータが、ディスクドライブ14a~14dのどのディスクドライブのどの位置に存在するかを記録したデータテーブルの形式の例である。

#### 【0046】

データブロック番号、ディスクドライブ番号、ドライブ内データ位置、さらにはデータを送信した先の入出力チャネルを、送信先入出力チャネル番号として、これらを組みとして記録する。また未だ当該データブロック番号のデータをどの入出力チャネルにも送信したことがなければ、この送信先入出力チャネル番号としては、予め定める無効な入出力チャネル番号を記録しておく。あるいは、送信した入出力チャネルのキャッシュメモリが溢れて当該データが追い出され削除されたときにも、この送信先入出力番号として、無効な入出力チャネル番号を設定してもよい。

#### 【0047】

このデータテーブルは、それぞれのディスクアダプタ13a~13dにおいて、ディスクドライブ14a~14dのうちのそれぞれのディスクアダプタ13a~13dに接続されたディスクドライブに格納されているデータブロック番号の分だけ持てばよい。なお、このデータテーブルは、実際には例えば図3におけるローカルメモリ302に格納され、プロセッサ301によって必要に応じ読み書きされる。

#### 【0048】

### <コマンド形式>

キャッシュメモリに記憶しているデータに関する整合性確保処理において用い

られる無効化コマンドと更新コマンドの、それぞれのコマンド形式について説明する。

#### 【 0 0 4 9 】

図 8 は、ディスクアダプタ 13a～13d において、入出力チャネル 10a～10d のうちの一つに一旦送信したデータを他の入出力チャネル 10a～10d から要求されたときに、先に送った方の入出力チャネル 10a～10d のキャッシュメモリ 11a～11d 内のデータを無効化するために当該入出力チャネル 10a～10d に向けて送信する無効化コマンドのコマンド形式の例である。無効化コマンドは、無効化コマンドコードと、無効化されるデータブロック番号からなる。

#### 【 0 0 5 0 】

図 9 は、同一データを持つ複数の入出力チャネル 10a～10d 間である入出力チャネル 10a～10d のデータの書き換えが起こった場合に、これを当該ディスクアダプタ 13a～13d を通じてディスクドライブ 14a～14d に書き戻すと同時に、他の入出力チャネル 10a～10d のキャッシュメモリ 11a～11d のデータも書き換えて更新する場合に、図 8 の無効化コマンドに替えて、ディスクアダプタ 13a～13d より入出力チャネル 10a～10d に送信する更新コマンドのコマンド形式の例である。更新コマンドを表すコマンドコード、当該データブロック番号に加えて、新しくデータ内容を更新するための更新データからなる。

#### 【 0 0 5 1 】

#### <処理説明>

つぎに、ホストコンピュータから送られてくる、データ要求やデータ書き込み要求などのデータ入出力要求を受信した場合に行われるディスクアレイ装置 1 の処理について詳述する。

#### 【 0 0 5 2 】

図 1 0 は、ホストコンピュータからディスクアレイ装置 1 に送られてくるデータ入出力要求のデータフォーマットの一例を示す。ヘッダにはホストコンピュータの ID などが設定される。ポート ID には、そのデータ入出力要求の処理対象となる論理ボリュームが接続するポートのポート ID が設定される。ここで論理ボリュームとは、ディスクドライブ 14a～14d の記憶領域上に区画設定された論理

的な記憶領域である。LUN (Logical Unit Number) の欄には、当該データ入出力要求が対象とする論理ボリュームの識別子が設定される。アドレスの欄およびデータ長の欄は、例えば、データ入出力要求がデータ要求である場合に設定され、この場合、アドレスの欄には読み出し先のデータが格納されている論理ボリュームの先頭アドレスが設定され、また、データ長の欄には読み出し対象となるデータのデータ長が設定される。

## 【 0 0 5 3 】

モードの欄には、そのデータ入出力要求に対するホストコンピュータへの応答処理と、各キャッシュメモリのデータの整合性確保処理の実行順を制御するためにディスクアレイ装置1に参照される値が設定される。この実施例では、モードの欄には、「0」または「1」のいずれかの値が設定されるものとする。

## 【 0 0 5 4 】

なお、「0」は、例えば、そのデータ入出力要求の処理について、各キャッシュメモリに記憶しているデータの整合性を確保することよりも、ホストコンピュータに対する応答速度が要求される場合に設定される。また、「1」は、例えば、そのデータ入出力要求の処理について、各キャッシュメモリに記憶しているデータの整合性を確保する必要がある場合に設定される。

## 【 0 0 5 5 】

## &lt;データ要求&gt;

まず、ホストコンピュータから送信されてくるデータ入出力要求として、データ要求を受信した場合にディスクアレイ装置1において行われる処理を、図11に示すフローチャートとともに説明する。なお、ここではデータ要求を入出力チャンネル10aが受信した場合について説明するが、他の入出力チャンネル10b～10dが受信した場合でも同様である。

## 【 0 0 5 6 】

ホストコンピュータからデータ要求を受信 (1110) した入出力チャンネル10aは、まず、読み出し対象となるデータがキャッシュメモリ11a上に存在するかどうかを調べる (1112)。ここで存在する場合 (1112: YES) には、入出力チャンネル10aはそのデータをキャッシュメモリ11aから読み出して、これをホストコンピュー

タに送信する (1113)。

【 0 0 5 7 】

一方、キャッシュメモリ11aに要求されたデータが存在しない場合 (1112:NO) には、入出力チャネル10aはデータ要求で指定されているデータブロック番号に基づいて当該データが存在するディスクドライブに接続されているディスクアダプタ13a~13dを決定し、データ要求をネットワークスイッチ12aまたは12bを用いてそのディスクアダプタ13a~13dに送信する。ディスクアダプタ13aは、データ要求を受け取ると、そのデータブロック番号に基づいて具体的にデータが存在するディスクドライブ14a~14d内のディスクドライブを決定し (1114)、そのディスクドライブからのデータの読み出しを開始する (1115)。

【 0 0 5 8 】

ここでデータ要求のモードの欄の内容を調べ (1116)、モード欄に「0」が設定されていた場合 (1116:YES) には、ディスクドライブから読み出したデータを入出力チャネル10aに送信 (1117) した後、(1118) 以下の処理に進む。

一方、モード欄に「1」が設定されていた場合 (1116:NO) には、そのまま (1118) 以下の処理に進む。

【 0 0 5 9 】

(1118) では、ディスクアダプタ13aが、ここで要求されているデータを他の入出力チャネル10aとは別の入出力チャネル10b~10dに以前に送信したことがあるかどうかを検索する。この検索は、具体的には、ディスクアダプタ13aにおいて、図7のデータテーブルの送信先入出力チャネル番号に、有効な入出力チャネルの番号が登録されているかどうかを検索することにより行われる。

【 0 0 6 0 】

ここで他の入出力チャネルのキャッシュメモリに送ったことがない場合には、(1125) へ進む。一方、他の入出力チャネルのキャッシュメモリに送ったことがある場合には、その入出力チャネルに図8に示した無効化コマンドを送信し (1120)、当該データをその入出力チャネルのキャッシュから削除させる。

【 0 0 6 1 】

この際その入出力チャネルのキャッシュメモリにおいて当該データの書き換え

が生じてまだその書き換えられたデータがディスクドライブに書き戻していなかったかどうか判定される（1121）。この判定は、具体的には図4または図5における書き換えビットD 402または502を検査することにより行われる。書き戻していない場合には、その入出力チャネルにおいて書き戻していないデータをそのキャッシュメモリから読み出し、当該ディスクアダプタに返送させる。そして、当該ディスクアダプタは、（1115）で開始したディスクドライブからのデータの読み出しが完了するのを待って、ディスクドライブから読み出されたデータの上に、入出力チャネルから返送されてきた書き換えデータを上書きする（1122）。

## 【 0 0 6 2 】

つぎにディスクアダプタは、その入出力チャネルから無効化処理完了信号が送信されてくるのを待ち、無効化されたことの確認を行う（1123）。

## 【 0 0 6 3 】

つぎに図7に示したデータテーブルにおいて、送信先入出力チャネル番号に予め定める無効なチャネル番号を設定することで、データ送信先の記録を削除する（1124）。なお、この後この送信先入出力チャネル番号には、後述する（1125）で新たにデータを送信した入出力チャネル番号を記録するので、この（1124）の処理はなくてもよい。

## 【 0 0 6 4 】

つぎに、（1110）新たにデータを要求してきた入出力チャネルのキャッシュメモリにデータを送信することを決定し、この入出力チャネルを記録する。具体的には、この記録は、図7のようなデータテーブルの送信先入出力チャネル番号に有効なチャネル番号で記録することにより行われる。

## 【 0 0 6 5 】

つぎに、もしまだ（1115）で開始したディスクドライブの読み出しが完了していなかった場合には、読み出しが完了するまで待ち、読み出したデータを当該入出力チャネルに送信して、その入出力チャネルに接続されたキャッシュメモリに格納させる（1126）。

## 【 0 0 6 6 】

つぎにディスクアダプタは、データ要求のモードの欄の内容を調べ（1127）、



モード欄に「1」が設定されていた場合（1127:YES）には、ディスクドライブから読み出したデータを入出力チャネル10aに送信（1128）した後、（1129）以下の処理に進む。一方、モード欄に「0」が設定されていた場合（1127:NO）には、そのまま（1129）以下の処理に進む。

【 0 0 6 7 】

（1129）では、（1122）を通過し、ディスクドライブから読み出したデータに対して上書きが生じたかどうかを判定する。もし上書きが生じた場合には、上書きされたデータを当該ディスクドライブへ書き戻す（1130）。

【 0 0 6 8 】

なお、以上の処理において、モード欄の設定値が「1」であった場合、すなわち、ホストコンピュータに応答を返すより前に無効化処理を行う場合には、ディスクドライブからのデータの読み出し（1115）の処理を、（1122）や（1126）の処理完了までに完了することにより、すなわち、これらの処理をオーバーラップさせて処理するようにし、処理速度の向上を図るようにしてもよい。

【 0 0 6 9 】

以上に説明したように、この実施例では、（1116）において、モード欄に「0」が設定されている場合には、ディスクドライブから読み出したデータを先にホストコンピュータに送信した後に、（1118）に後続する処理を実行する。すなわち、そのデータ入出力要求のモード欄に「0」が設定されている場合には、各キャッシュメモリに記憶しているデータに関する整合性を確保するために必要となる（1118）～（1130）の処理は、ディスクアレイ装置1からホストコンピュータへの応答処理が行われた後で行われることとなり、これによりホストコンピュータに対する応答速度が向上する。

【 0 0 7 0 】

従って、ホストコンピュータで動作するソフトウェアの仕様や性質などにより、整合性を確保する必要がなかったり、また、必要があるとしてもさほど厳密性が要求されないような場合には、そのようなソフトウェアからのデータ入出力要求のモード欄に「0」を設定しておくことで、このようなデータ入出力要求についてはホストコンピュータに対する応答速度を向上させることができる。

## 【 0 0 7 1 】

とくに、(1120)などの無効化に関する処理では、無効化要求や無効化がされたことの確認のための通信がスイッチを介して行われ、ホストコンピュータへの応答速度を低下させる要因となるが、このように無効化処理が事後的に行われることで、ホストコンピュータに対する応答速度が改善されることになる。

## 【 0 0 7 2 】

なお、図 1 1 では、(1118)の前にモード「0」の判定とホストコンピュータへの送信処理(1116,1117)を行うようにしているが、これらの処理は、例えば(1118)と(1119)の間や(1119)と(1120)の間で行っても良い。つまり、各キャッシュメモリに記憶しているデータに関する整合性を確保するために必要となる処理のうち、ホストコンピュータへの応答速度を低下させる主因となる処理が行われるよりも前に、ホストコンピュータに応答を返す構成であれば、ホストコンピュータへの応答速度の向上が期待できることになる。

## 【 0 0 7 3 】

## ＜データ書き込み要求＞

つぎに、ホストコンピュータから送信されてくるデータ入出力要求として、データ書き込み要求を受信した場合にディスクアレイ装置1において行われる処理について説明する。なお、ここではデータ書き込み要求を入出力チャネル10aが受信した場合について説明するが、他の入出力チャネル10b～10dが受信した場合でも同様である。

## 【 0 0 7 4 】

図 1 2 に示すフローチャートは、ディスクアダプタ13a～13dが、入出力チャネル10a～10dのうちの一つに一旦送信したデータを他の入出力チャネル10a～10dから要求されたときに、先に送った方の入出力チャネル10a～10dのキャッシュメモリ11a～11d内のデータを無効化するための無効化信号を該当の入出力チャネル10a～10dに向けて送信することで各キャッシュメモリ11a～11dのデータに関する整合性を確保するようにしている場合に、この整合性確保の処理とホストコンピュータへの完了通知との実行順を、データ入出力要求の内容であるモードに応じて制御するようにした場合のディスクアレイ装置1の処理についての実施例を説明

している。

【 0 0 7 5 】

ホストコンピュータからデータ書き込み要求を受信 (1211) した入出力チャネル10aは、そのデータ書き込み要求のモードの欄の内容を調べる (1212)。ここでモードの欄に「0」が設定されていた場合 (1212: YES)、当該入出力チャネル10aのキャッシュメモリ11aへの書き込みを許可して書き込みデータをキャッシュメモリ11aに書き込み (1213)、書き込み完了通知をホストコンピュータに送信する (1214)。一方、(1212) において、モード欄に「1」が設定されていた場合 (1212: NO) には、入出力チャネル10aはすぐにキャッシュメモリ11aにデータを書き込まず、(1215) に進む。

【 0 0 7 6 】

つぎに、入出力チャネル10aは、書き込み要求があったことを該当のディスクアダプタ13a~13dに通知し、通知を受けたディスクアダプタは、書き込み要求があった入出力チャネル10a以外に、当該データを送信した入出力チャネルのキャッシュメモリがあったかどうかを検索する (1215)。ここでもし他に送信した入出力チャネルのキャッシュメモリがなければ (1216: NO)、(1220) に進む。一方、他に送った入出力チャネルのキャッシュメモリがあれば、それらの入出力チャネルのすべてのキャッシュメモリを対象として該当の入出力チャネルに無効化要求を送信する (1217)。ここでこの無効化要求には、例えば図8に示した、無効化コマンドを用いることができる。

無効化されたことが確認されたら (1218)、無効化したキャッシュメモリの入出力チャネル番号を、送信先入出力チャネル番号の記録から削除する (1219)。

【 0 0 7 7 】

つぎに (1220) では、そのデータ書き込み要求のモード欄の内容を調べる。ここでモード欄に「1」が設定されている場合には、(1221) に進み、書き込み要求を受け取った入出力チャネル10aのキャッシュメモリ11aへの書き込みを許可し、ホストコンピュータからの書き込みデータを受信してキャッシュメモリ11aへの書き込みを行い (1221)、また、書き込み完了通知をホストコンピュータに送信する (1222)。そして、キャッシュメモリ11aに書き込んだデータを当該入出

力チャンネル10aより該当のディスクアダプタ13a～13dに送信し、ディスクドライブ14a～14dにも書き戻しを行わせる（1223）。

他方、（1220）においてモード欄に「0」が設定されていた場合には（1221）に相当する処理は既に（1213）にて行われているので（1223）に進む。

#### 【0078】

以上に説明したように、この実施例のディスクアレイ装置1は、モード欄に「0」が設定されている場合には、ディスクドライブ14a～14dから読み出したデータを、先にホストコンピュータに送信してしまった後に、整合性確保に関する処理を実行する。すなわち、そのデータ入出力要求にモードが「0」が設定されている場合には、整合性確保に関する処理は、ディスクアレイ装置1からホストコンピュータへの応答処理が行われた後に事後的に行われることとなり、ホストコンピュータに対する応答速度の向上が図られる。

#### 【0079】

従って、ホストコンピュータで動作するソフトウェアの仕様や性質などにより、整合性を確保する必要が無かったり、また、必要があるとしても厳密性が要求されないような場合には、そのようなソフトウェアからのデータ入出力要求のモード欄には「0」を設定しておけば、このようなデータ入出力要求についてはホストコンピュータに対する応答速度を向上させることができる。

#### 【0080】

図13に示すフローチャートは、ある入出力チャンネル10a～10dに対応するキャッシュメモリ11a～11dのデータの書き換えが生じた場合に、これをディスクアダプタ13a～13dを通じてディスクドライブ14a～14dに書き戻すと同時に、他の入出力チャンネル10a～10dのキャッシュメモリ11a～11dのデータを書き換えて更新することで各キャッシュメモリ11a～11dのデータに関する整合性を確保するようにしている場合に、この整合性確保の処理とホストコンピュータへの完了通知との実行順を、データ入出力要求の内容であるモードに応じて制御するようにした場合のディスクアレイ装置1の処理についての実施例を説明している。

#### 【0081】

ホストコンピュータからデータ書き込み要求を受信（1311）した入出力チャネ

ル10aは、当該入出力チャネル10aのキャッシュメモリ11aへの書き込みを許可して書き込みデータをキャッシュメモリ11aに書き込む（1312）。

【 0 0 8 2 】

つぎに、そのデータ書き込み要求のモードの欄の内容を調べ（1313）、モードの欄に「0」が設定されていた場合（1313:YES）には、書き込み完了通知をホストコンピュータに送信する（1314）。一方、（1313）において、モード欄に「1」が設定されていた場合（1313:NO）には（1315）に進む。

【 0 0 8 3 】

（1315）では、適宜なタイミングで当該入出力チャネル10aから書き込みデータを該当のディスクアダプタ13a～13dに送信し、キャッシュメモリ11aに書き込んだデータを元のディスクドライブ14a～14dにも書き戻す（1315）。

【 0 0 8 4 】

つぎに、ディスクアダプタ13a～13dは、書き込みが行われた入出力チャネル10a以外に、以前にそのデータを送信した入出力チャネルのキャッシュメモリがあったかどうかを検索する（1316）。もし他に送った入出力チャネルのキャッシュメモリがあれば（1317:YES）、（1318）に進む。無ければ（1319）に進む。

【 0 0 8 5 】

（1318）では、当該データをもつ他の入出力チャネルのキャッシュメモリのすべてに当該ディスクアダプタからディスクドライブに書き込んだ書き込みデータを送信し、データの変更を行わせる。

【 0 0 8 6 】

（1319）では、データ書き込み要求のモードの欄の値を調べ、モードの欄に「1」が設定されていた場合（1319:YES）には、書き込み完了通知をホストコンピュータに送信する（1320）。一方、モードの欄に「0」が設定されていた場合（1320:NO）には、完了報告は既に（1314）で行っているのでここでは行わない。

【 0 0 8 7 】

以上のように、この実施例では、モード欄に「0」が設定されている場合には、ディスクドライブ14a～14dから読み出したデータを、先にホストコンピュータに送信してしまった後に、整合性確保に関する処理を実行する。すなわち、モー

ドの欄に「0」が設定されている場合には、整合性確保に関する処理は、ディスクアレイ装置1からホストコンピュータへの応答処理が行われた後で行われ、整合性確保に関する処理よりも前にホストコンピュータへの応答がされることになるので、ホストコンピュータへの応答速度が向上することになる。

## 【0088】

このため、ホストコンピュータで動作するソフトウェアの仕様や性質などにより、各キャッシュメモリ間でのデータに関する整合性を確保する必要が無かったり、また、必要があるとしてもさほど厳密性が要求されないような場合には、そのようなソフトウェアからのデータ入出力要求のモード欄に「0」を設定しておくことで、ホストコンピュータに対する応答速度を向上させることができる。

## 【0089】

## ＜他の実施形態＞

前述の実施例では、データ入出力要求のモード欄の値に応じて、そのデータ入出力要求に対するホストコンピュータへの応答処理と、各キャッシュメモリ11a～11dのデータに関する整合性確保処理の実行順を制御するようにしているが、モード欄に代えて、データ入出力要求に記述されているLUNの値に応じて処理の切り替えを行うようにすることもできる。

## 【0090】

具体的には、例えば、図11～図13に示したフローチャートにおけるモード欄の値により実行制御を行っている部分において、モード欄の値に代えて、LUNを参照するようにし、LUNの値に応じて処理を分岐させるようにする。なお、モード欄の設定値やLUNごとに前記実行順をどのように制御するかを、例えば、オペレータがディスクアレイ装置1の管理用端末を操作して設定できるようにすることもできる。

## 【0091】

また、最近のディスクアレイ装置の中には、アベイラビリティの向上や、ディザスタリカバリ性能の向上などを目的として、ディスクアレイ装置がデータ入出力要求に応じて自身の論理ボリュームに管理しているデータの内容を、遠隔地などに設置され当該ディスクアレイ装置に接続する他のディスクアレイ装置にも記

憶する機能（「リモートコピー機能」と称されることもある）を備えるものが存在する。

【 0 0 9 2 】

そして、この機能の実現方式として、ディスクアレイ装置1が備える入出力チャネルに他のディスクアレイ装置を通信可能に接続し、各入出力チャネルを自身のディスクドライブに書き込ませる動作モード（以下、「ノーマルモード」と称する）で動作させるか、他のディスクアレイ装置にデータ入出力要求を送信させる動作モード（以下、「コピーモード」と称する）で動作させるかをホストコンピュータからディスクアレイ装置1に送信するコマンドにより切り替えられるようにし、ホストコンピュータから入出力チャネルの動作モードを切り替えるようにしていることがある。

【 0 0 9 3 】

ここでこのような方式で前記機能が実現されている場合には、データ入出力要求を受信した入出力チャネルが、いずれの動作モードで動作しているかに応じて、そのデータ入出力要求に対するホストコンピュータへの応答処理と、各キャッシュメモリのデータに関する整合性確保処理の実行順を制御するようにすることが考えられる。すなわち、例えば、コピーモードで動作している入出力チャネルで受信したデータ入出力要求については、ホストコンピュータへの応答を優先させて、整合性確保のための処理は後で行うようにし、また、ノーマルモードで動作している入出力チャネルで受信したデータ入出力要求については、整合性確保のための処理を優先させ、その後でホストコンピュータへの応答を行うようにする。

【 0 0 9 4 】

つまり、コピーモードで動作中の入出力チャネルで受信したデータ入出力要求の処理は、ノーマルモードで動作中の入出力チャネルで受信したデータ入出力要求の処理に比べて整合性確保の必要性が少ないことが多いと考えられ、このような場合には、この仕組みを適用することでホストコンピュータへの応答速度を向上させることができる。

【 0 0 9 5 】

以上、本発明に係るディスクアレイ装置等について説明してきたが、上記した発明の実施の形態は、本発明の理解を容易にするためのものであり、本発明を限定するものではない。本発明は、その趣旨を逸脱することなく、変更、改良され得ると共に、本発明にはその等価物が含まれることは勿論である。

【 0 0 9 6 】

【発明の効果】

本発明によれば、ディスクアレイ装置、ディスクアレイ装置の制御方法、及びストレージシステムを提供することができる。

【図面の簡単な説明】

【図 1】 本発明の一実施例による、ストレージシステムの概略構成を示す図である。

【図 2】 本発明の一実施例による入出力チャネルの構成を示す図である。

【図 3】 本発明の一実施例によるディスクアダプタの構成を示す図である。

【図 4】 本発明の一実施例によるキャッシュメモリの構成を示す図である。

【図 5】 本発明の一実施例によるキャッシュメモリの構成を示す図である。

【図 6】 本発明の一実施例によるデータテーブルの形式を示す図である。

【図 7】 本発明の一実施例によるデータテーブルの形式を示す図である。

【図 8】 本発明の一実施例による無効化コマンドのコマンド形式を示す図である。

【図 9】 本発明の一実施例による、更新コマンドのコマンド形式を示す図である。

【図 1 0】 本発明の一実施例による、データ入出力要求のデータフォーマットを示す図である。

【図 1 1】 本発明の一実施例による、データ要求を受信した場合にディスクアレイ装置において行われる処理を説明するフローチャートを示す図である。

【図 1 2】 本発明の一実施例による、データ書き込み要求を受信した場合にディスクアレイ装置において行われる処理を説明するフローチャートを示す図である。

【図 1 3】 本発明の一実施例による、データ書き込み要求を受信した場合に



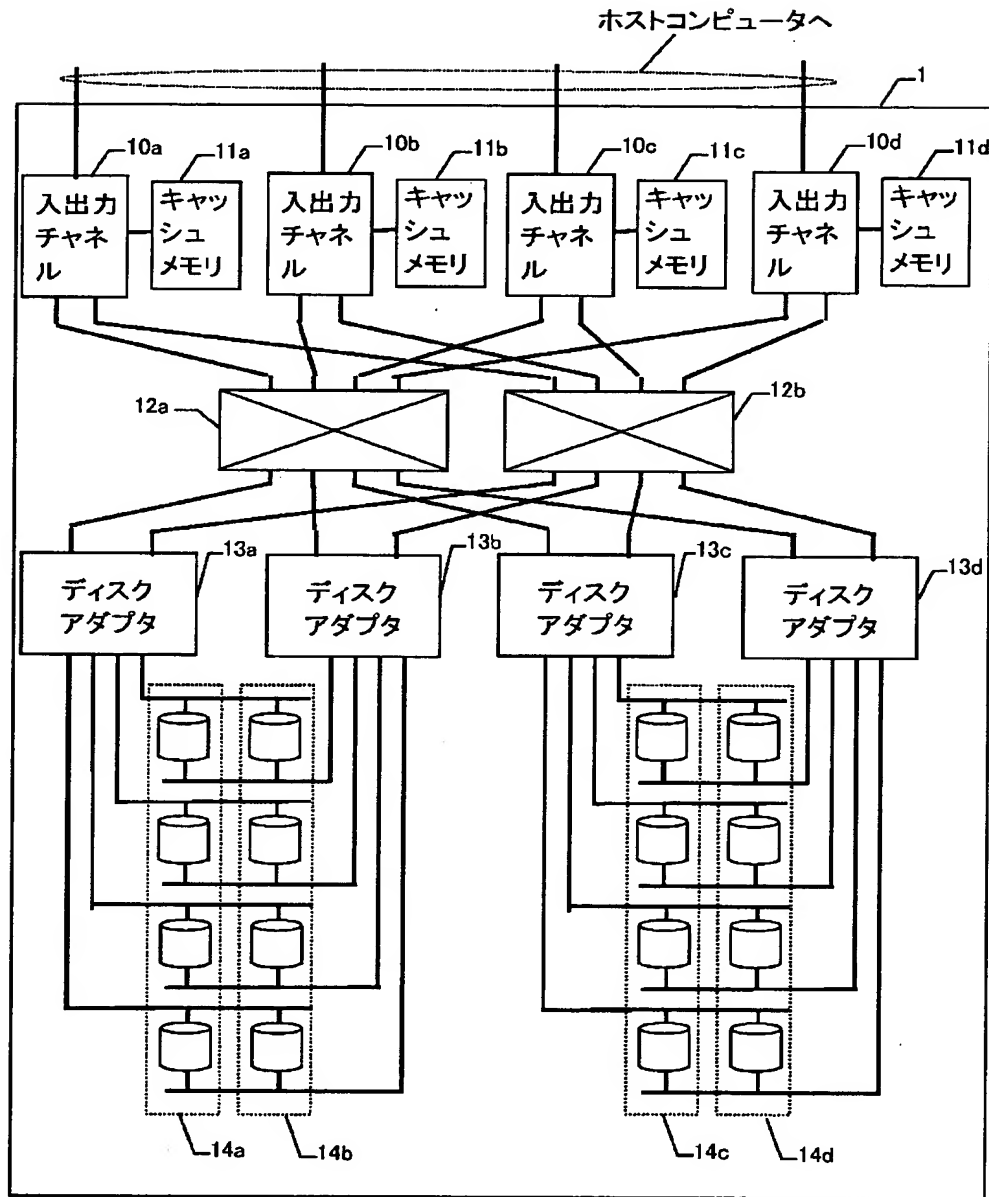
ディスクアレイ装置において行われる処理を説明するフローチャートを示す図である。

【符号の説明】

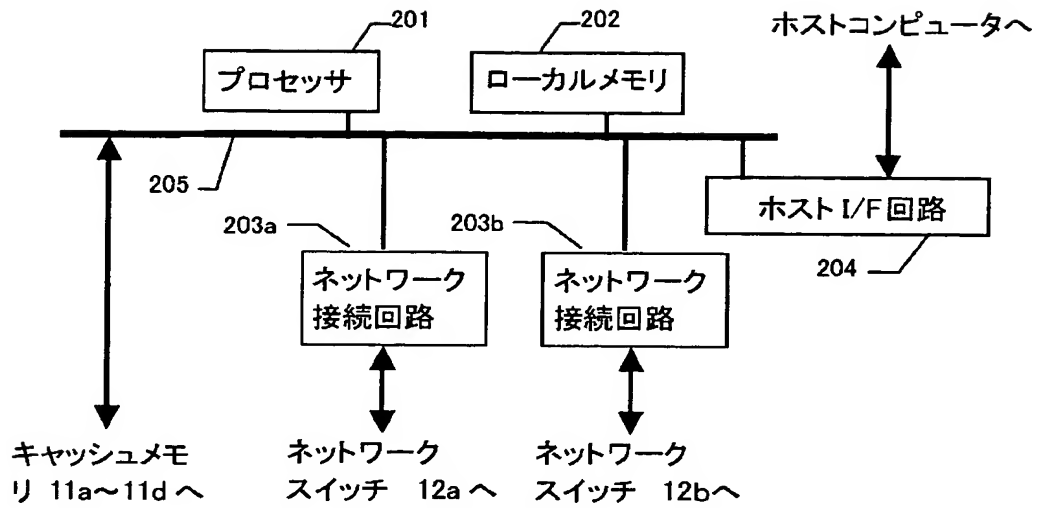
- 1 ディスクアレイ装置
  - 1 0 a ~ 1 0 d 入出力チャネル
  - 1 1 a ~ 1 1 d キャッシュメモリ
  - 1 2 a、1 2 b ネットワークスイッチ
  - 1 3 a ~ 1 3 d ディスクアダプタ
  - 1 4 a ~ 1 4 d ドライブ群
- 2 0 1、3 0 1 プロセッサ
- 2 0 2、3 0 2 ローカルメモリ
- 2 0 3 a、2 0 3 b、3 0 3 a、3 0 3 b ネットワーク接続回路
- 2 0 4 ホスト I / F 回路
- 3 0 4 ディスクドライブ I / F 回路
- 3 0 5 プロセッサバス
- 4 0、5 0 検索制御部
- 4 1 メモリ部
  - 4 1 1、5 1 1 アドレスバス
  - 4 1 2、5 1 2 返答信号バス
  - 4 1 3、5 1 3 データバス
  - 4 1 4、5 1 4 制御コマンドバス
- 5 1 タグ部
- 5 2 データ部
- 5 1 5 データ位置信号

【書類名】 図面

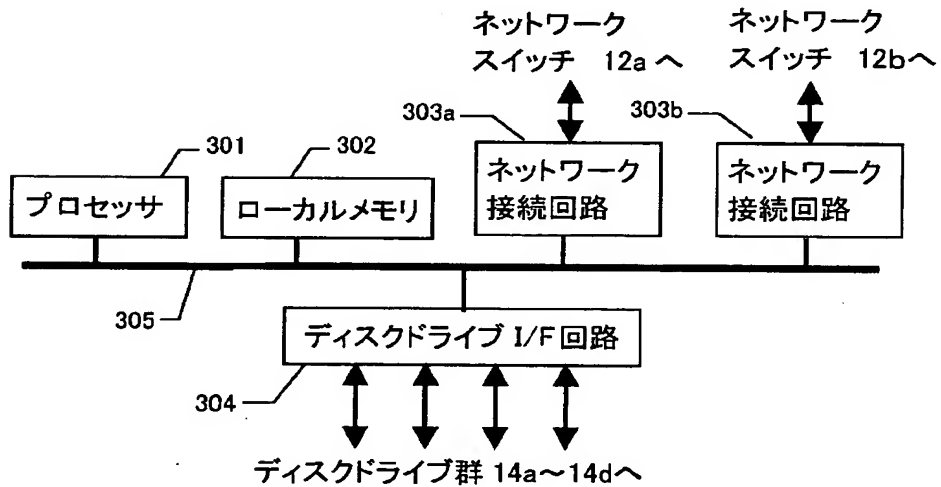
【図 1】



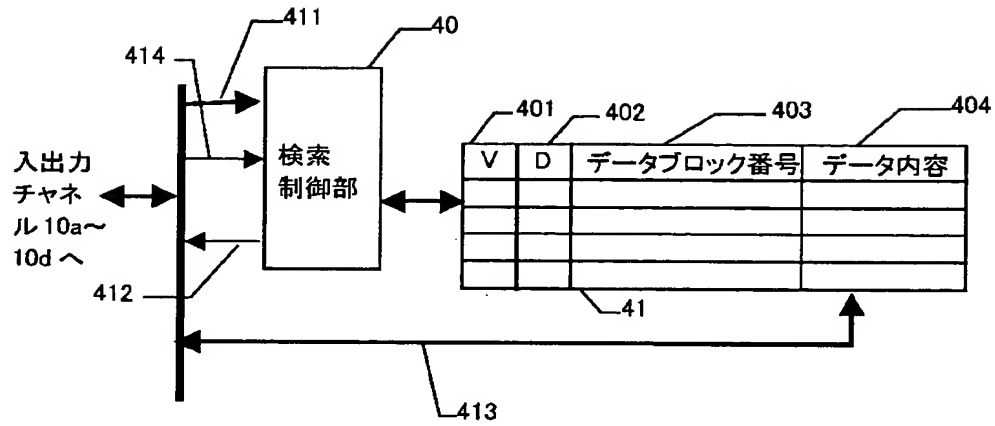
【図 2】



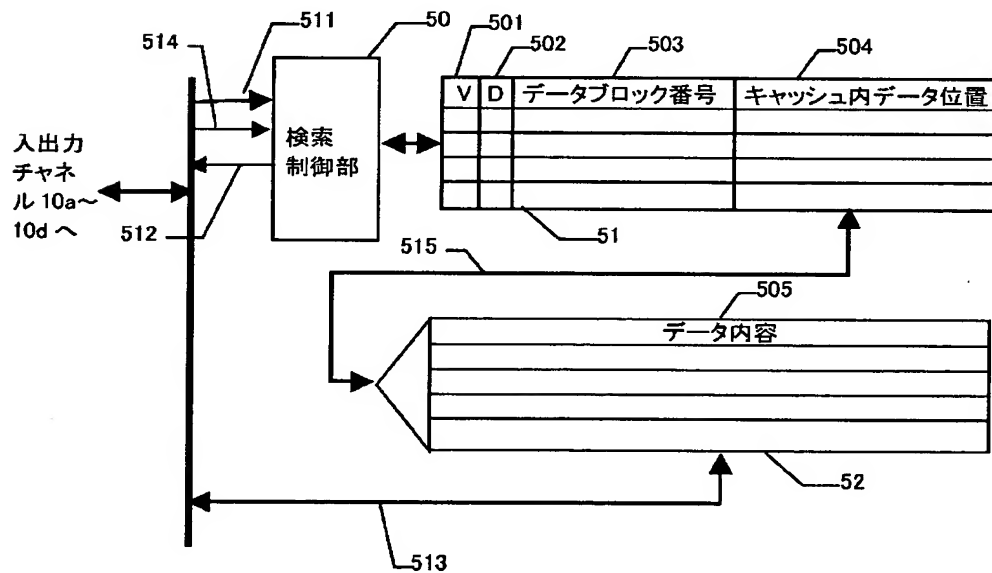
【図 3】



【図 4】



【図 5】



【図 6】

データブロック番号	ディスクアダプタ番号

【図 7】

データブロック番号	ディスクドライブ番号	ドライブ内データ位置	送出先入出力チャンネル番号

【図 8】

無効化コマンドコード	データブロック番号
------------	-----------

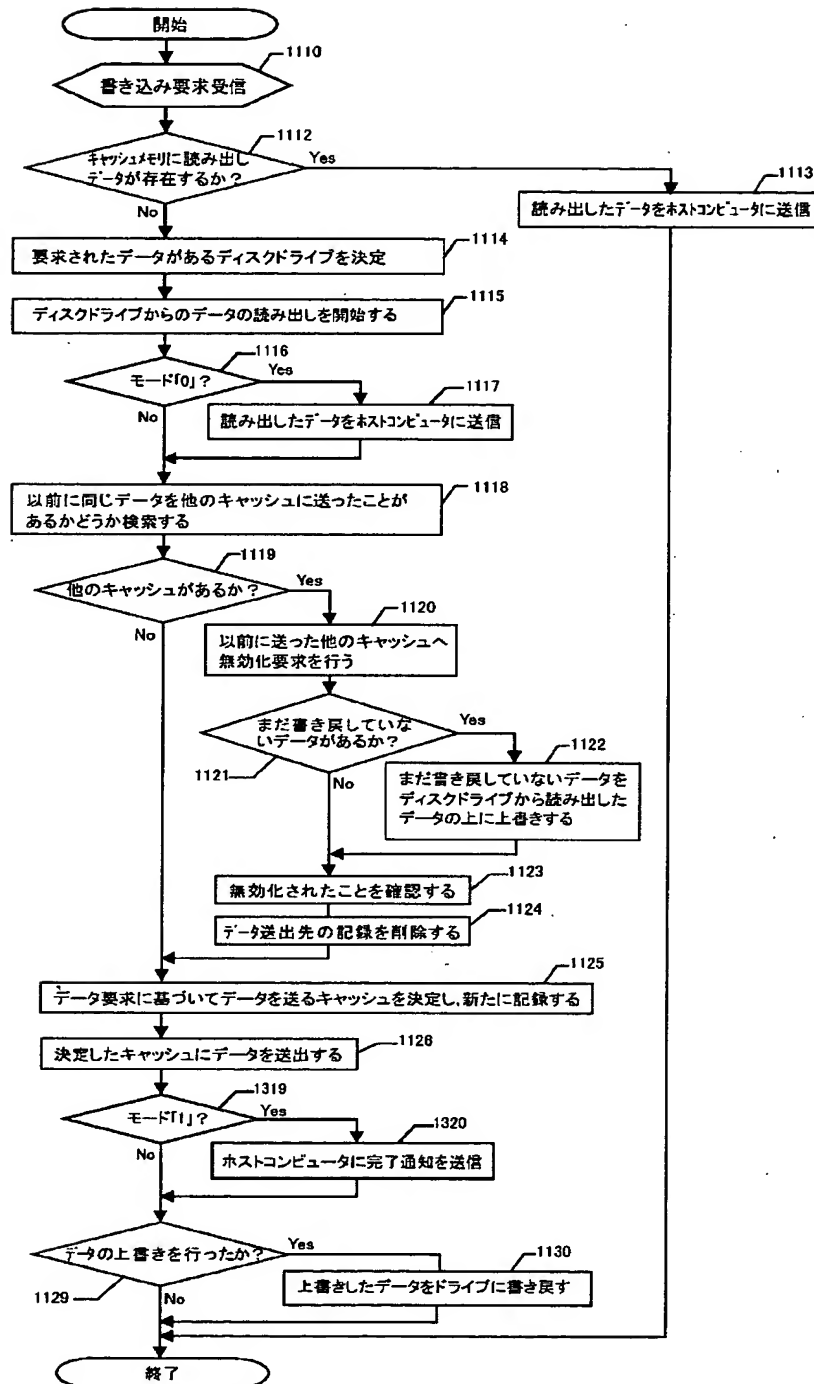
【図 9】

更新コマンドコード	データブロック番号	更新データ
-----------	-----------	-------

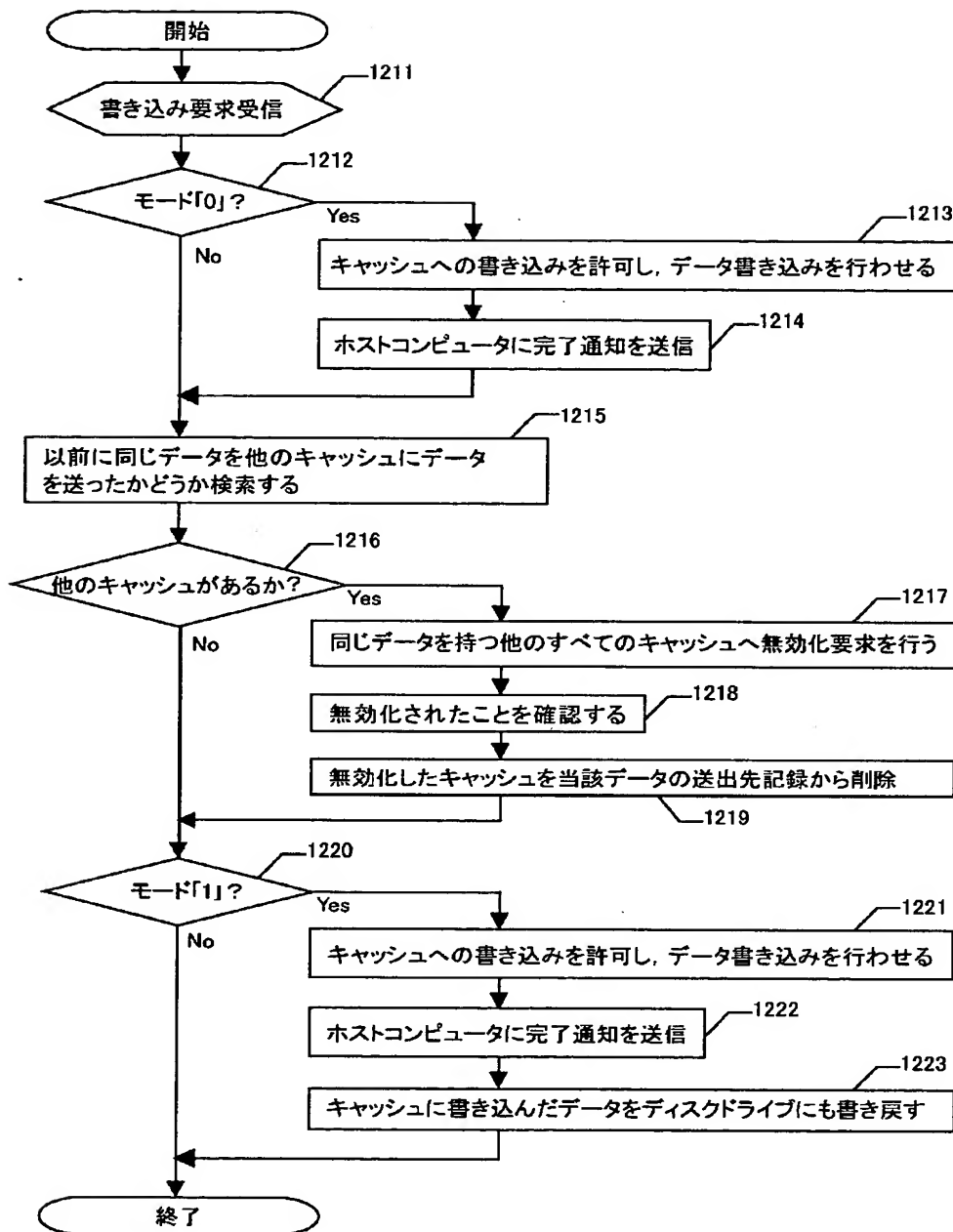
【図 1 0】

ヘッダ	ポートID	LUN	アドレス	データ長	モード
-----	-------	-----	------	------	-----

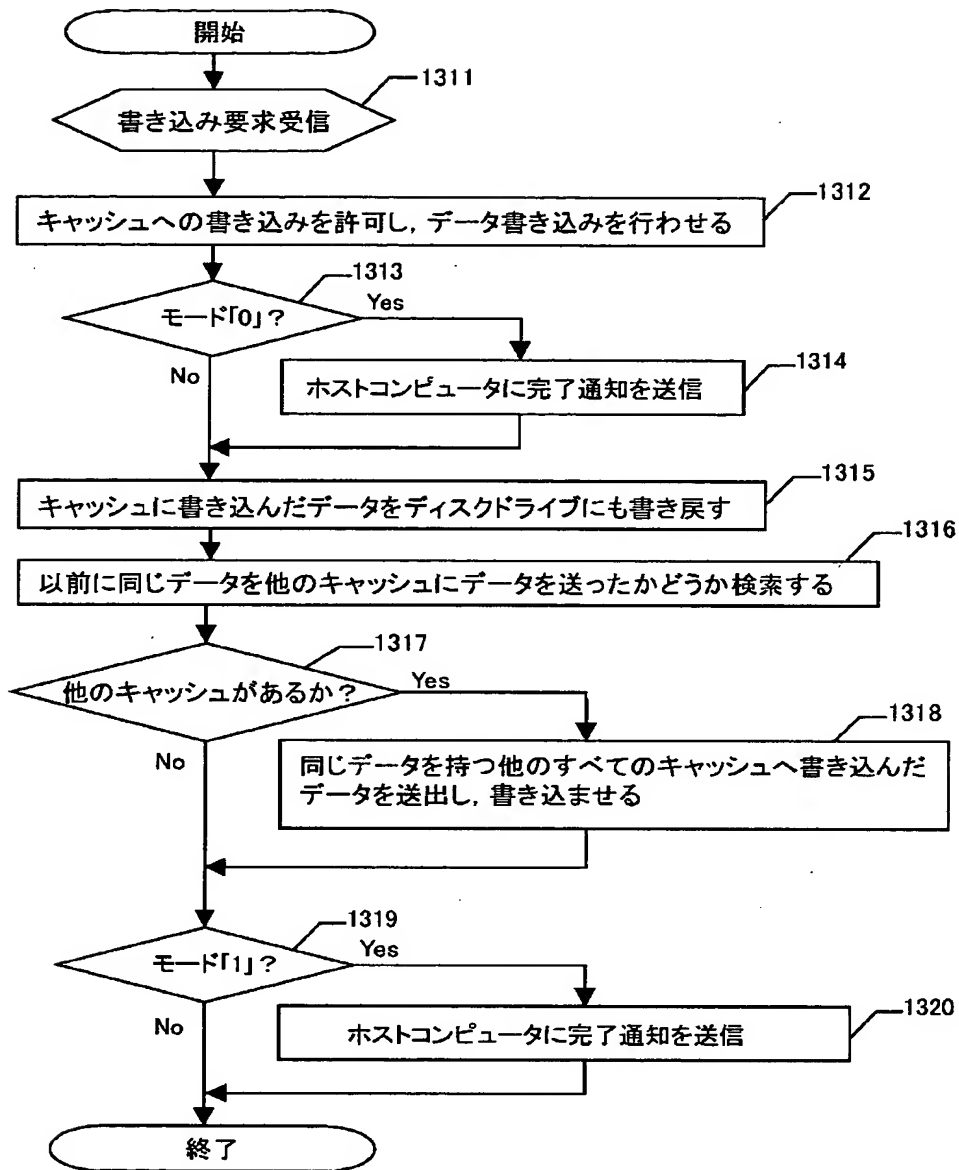
【図 11】



【図 12】



【図 1 3】





【書類名】 要約書

【要約】

【解決手段】 外部装置からのデータ入出力要求を受信する複数の入出力チャンネルと、入出力チャンネルごとに用意されそれぞれ対応する入出力チャンネルに接続する複数のキャッシュメモリと、ディスクドライブと、ディスクドライブに対してデータ入出力を行うディスク制御手段と、入出力チャンネルとディスク制御手段とを通信可能に接続する通信手段と、を備えるディスクアレイ装置を提供する。

【選択図】 図 1 1

出 願 人 履 歴 情 報

識別番号 [ 0 0 0 0 0 5 1 0 8 ]

1. 変更年月日 1 9 9 0 年 8 月 3 1 日

[ 変更理由 ] 新規登録

住 所 東京都千代田区神田駿河台 4 丁目 6 番地  
氏 名 株式会社日立製作所